CONCEPT

SQL for Data Analysis: Querying, Cleaning and Preparing Data

22 – 26 March 2027
London - Central London four-star



SQL for Data Analysis: Querying, Cleaning and Preparing Data

Ref: 1247_7391 **Date:** 22 – 26 March 2027 **Location:** London - Central London four-star **Fees:** 25300 SAR

Technical depth: Practitioner · **Practical mode:** Lab

Introduction

Analysts often spend most of their time exporting data to spreadsheets, fixing it by hand and repeating the same steps every reporting cycle, which introduces errors and leaves no audit trail. Querying the source directly with SQL makes preparation faster, repeatable and checkable. This Core Concept course equips analysts to query relational databases, clean and reshape data and validate results using standard SQL that works across common database platforms. Participants work hands on with datasets from several sectors and leave with a SQL Data Preparation Script Pack and data dictionary for a reporting need of their own.

Course Objectives

- Retrieve, filter and sort data from relational databases using standard SQL in a query tool such as DBeaver or SQL Server Management Studio
- Combine tables correctly with joins and aggregations based on an understanding of keys, normalisation and star schemas
- Profile and clean text, numeric and date fields and recode values into analysis-ready formats
- Apply common table expressions, window functions and pivoting to answer multi-step analytical questions
- Detect duplicates, NULL errors and double counting and reconcile query results to control totals
- Produce a documented SQL Data Preparation Script Pack that another analyst can rerun and review

Target Audience

- Data and business analysts preparing datasets for reports and dashboards
- Reporting officers who currently extract and clean data manually in spreadsheets
- Finance, sales and operations analysts working with transactional systems
- Business intelligence support staff maintaining queries behind dashboards
- Research and statistics officers compiling data from administrative systems

Course Outline

Day 1: Relational Data and First Queries

- Relational Model: Tables, Primary Keys, Foreign Keys and Entity-Relationship Diagrams
- ISO/IEC 9075:2023 SQL Standard and Vendor Dialects in PostgreSQL and SQL Server
- Query Environment Set-Up in DBeaver and SQL Server Management Studio
- SELECT, WHERE and ORDER BY Filtering Fundamentals
- Source Data Inventory and Data Access Request Checklist

Day 2: Data Structures, Joins and Aggregation

- Normalisation from First to Third Normal Form
- Star Schema Fact and Dimension Tables in Kimball Modelling
- INNER, LEFT and FULL OUTER JOIN Logic
- Aggregation with GROUP BY, HAVING and Aggregate Functions
- CRISP-DM Data Understanding and Data Preparation Phases

Day 3: Profiling, Cleaning and Reshaping Data

- Data Profiling Queries: Row Counts, Distinct Values and NULL Rates
- Cleaning Text, Numeric and ISO 8601 Date Fields with SQL Functions
- CASE Expressions for Recoding, Banding and Flags
- Common Table Expressions for Readable Multi-Step Queries
- Tidy Data Principles and Reshaping with PIVOT and UNPIVOT

Day 4: Advanced Queries, Errors and Performance

- Window Functions: ROW_NUMBER, RANK, LAG and Running Totals
- Duplicate Detection and Resolution with Partitioned Queries
- Join Fan-Out, Three-Valued NULL Logic and Double-Counting Checks
- Execution Plans and Index Awareness for Query Performance
- Reusable Views, Parameterised Queries and SQL Injection Awareness

Day 5: Lab Practice and the Script Pack

- Sales and Customer Data Preparation Lab
- Public Service Request Log Preparation Lab
- Reconciliation Queries Against Source Control Totals
- SQL Data Preparation Script Pack and Data Dictionary Build
- Code Review Walkthrough with Git Version Control

Skills You Will Gain

- SQL Querying
- Relational Data Modelling
- Data Profiling
- Data Cleansing
- Analytical Window Functions
- Query Validation and Reconciliation
- Query Performance Awareness
- Script Documentation

Why Attend This Course

- Return to work with a SQL Data Preparation Script Pack and data dictionary for a real reporting need, already reviewed by peers
- Replace repetitive spreadsheet clean-up with scripts that rerun in minutes each reporting cycle
- Talk to database administrators and BI developers in their own terms when requesting data or access
- Compare data preparation practice with analysts from other sectors and countries working on similar systems

Conclusion

Reliable analysis starts with data that has been extracted, cleaned and combined correctly, in a way that can be repeated and checked. This course moves from the relational model and first queries, through table structures, joins and aggregation, to profiling, cleaning and reshaping data, and then to window functions, error traps and performance. The final day applies these skills to realistic datasets and produces a SQL Data Preparation Script Pack that participants take back to their own reporting work.